

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE Technical Report		3. DATES COVERED (From - To) -	
4. TITLE AND SUBTITLE Coordinated Beamforming for MISO Interference Channel: Complexity Analysis and Efficient Algorithms			5a. CONTRACT NUMBER W911NF-09-1-0279		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 611102		
6. AUTHORS Ya-Feng Liu, Yu-Hong Dai and Zhi-Quan Luo			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES University of Minnesota - Minneapolis Sponsored Projects Administration 450 McNamara Alumni Center Minneapolis, MN 55455 -2009			8. PERFORMING ORGANIZATION REPORT NUMBER 906.00		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 55069-MA.2		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT In a cellular wireless system, users located at cell edges often suffer significant out-of-cell interference. Assuming each base station is equipped with multiple antennas, we can model this scenario as a multiple-input single-output (MISO) interference channel. In this paper we consider a coordinated beamforming approach whereby multiple base stations jointly optimize their downlink beamforming vectors in order to simultaneously improve the data rates of a given group of cell edge users. Assuming perfect channel knowledge, we formulate this problem as the					
15. SUBJECT TERMS MISO interference channel, coordinated beamforming, complexity, cyclic coordinate descent algorithm, global convergence.					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Zhi-Quan (Tom) Luo
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 612-625-0242

Report Title

Coordinated Beamforming for MISO Interference Channel: Complexity Analysis and Efficient Algorithms

ABSTRACT

In a cellular wireless system, users located at cell edges often suffer significant out-of-cell interference. Assuming each base station is equipped with multiple antennas, we can model this scenario as a multiple-input single-output (MISO) interference channel. In this paper we consider a coordinated beamforming approach whereby multiple base stations jointly optimize their downlink beamforming vectors in order to simultaneously improve the data rates of a given group of cell edge users. Assuming perfect channel knowledge, we formulate this problem as the maximization of a system utility (which balances user fairness and average user rates), subject to individual power constraints at each base station. We show that, for the single carrier case and when the number of antennas at each base station is at least two, the optimal coordinated beamforming problem is NP-hard for both the harmonic mean utility and the proportional fairness utility. For general utilities, we propose a cyclic coordinate descent algorithm, which enables each transmitter to update its beamformer locally with limited information exchange, and establish its global convergence to a stationary point. We illustrate its effectiveness in computer simulations by using the space matched beamformer as a benchmark.

Coordinated Beamforming for MISO Interference Channel: Complexity Analysis and Efficient Algorithms*

Ya-Feng Liu[†], Yu-Hong Dai[†], and Zhi-Quan Luo[‡]

Abstract—In a cellular wireless system, users located at cell edges often suffer significant out-of-cell interference. Assuming each base station is equipped with multiple antennas, we can model this scenario as a multiple-input single-output (MISO) interference channel. In this paper we consider a coordinated beamforming approach whereby multiple base stations jointly optimize their downlink beamforming vectors in order to simultaneously improve the data rates of a given group of cell edge users. Assuming perfect channel knowledge, we formulate this problem as the maximization of a system utility (which balances user fairness and average user rates), subject to individual power constraints at each base station. We show that, for the single carrier case and when the number of antennas at each base station is at least two, the optimal coordinated beamforming problem is NP-hard for both the harmonic mean utility and the proportional fairness utility. For general utilities, we propose a cyclic coordinate descent algorithm, which enables each transmitter to update its beamformer locally with limited information exchange, and establish its global convergence to a stationary point. We illustrate its effectiveness in computer simulations by using the space matched beamformer as a benchmark.

Index Terms—MISO interference channel, coordinated beamforming, complexity, cyclic coordinate descent algorithm, global convergence.

I. INTRODUCTION

In a conventional wireless cellular system, base stations from different cells communicate with their respective remote terminals independently. Signal processing is performed on an intra-cell basis, while the out-of-cell interference is treated as background noise. This architecture often causes undesirable service outages to users situated near cell edges where the out-of-cell interference can be severe. Since the conventional intra-cell signal processing can not effectively mitigate the impact of inter-cell interference, we are led to consider coordinated base station beamforming across multiple cells in order to improve the services to edge users. In this paper, we focus on the downlink scenario where the base stations are equipped with multiple antennas and model it as a MISO interference

channel. We consider joint optimal beamforming across multiple base stations to simultaneously improve the data rates of a given group of cell edge users. Assuming that the channel state information (CSI) is known, we formulate this problem as the maximization of a system utility (which balances user fairness and average user rates), subject to individual power constraints at each base station. We show that, for the single carrier case and when the number of antennas at each base station is at least two, the optimal coordinated beamforming problem is NP-hard for both the harmonic mean utility and the proportional fairness utility. This NP-hardness result is in contrast to the single antenna case for which the same optimization problem is convex for both the harmonic mean and proportional fairness utility functions [1]. For the min-rate utility, this problem is known to be also solvable in polynomial time [1], [2].

In addition to the complexity analysis, we propose a practical iterative cyclic coordinate descent algorithm for the multi-cell coordinated beamforming problem by exploiting the separability of power constraints. We prove the global convergence of this cyclic coordinate descent algorithm (to a stationary point). Numerical experiments are also presented to illustrate the effectiveness of the proposed algorithm.

A. Related Work

Downlink beamforming has been studied extensively in the single cell setup [3], [4]. For the multi-cell interference channel, the reference [5] considered coordinated beamforming for the minimization of total weighted transmitted power across the base stations subject to individual signal-to-interference-plus-noise-ratio (SINR) constraints at the remote users. It turns out this problem can be transformed into a convex second order conic programming (SOCP) and efficiently solved. However, the maximization of weighted sum rates for a multi-cell interference channel under individual power constraints is NP-hard even for the single antenna and the single carrier case [1]. In fact, more is known about the single antenna interference channel case. For instance, if the system utility is changed into either the geometric mean rates (i.e., proportional fairness), the harmonic mean rates, or the min-rate, the corresponding utility maximization problem (for the single tone case) can be converted to a convex optimization problem and solved efficiently to global optimality [1], [6]. However, when the number of tones is more than two, then all of the aforementioned power control problems are NP-hard. The focus of this paper is to study the multi-antenna case (MISO interference

* Ya-Feng Liu and Yu-Hong Dai are supported by the National Natural Science Foundation of China, Grant number 10831006, and the CAS Grant kjcx-yw-s7-03. Zhi-Quan Luo is supported by the Army Research Office, Grant number W911NF-09-1-0279, and by the National Science Foundation, Grant number CMMI-0726336.

[†] State Key Lab. of Scientific and Engineering Computing, Beijing, 100190, China. (e-mail: {yafliu, dyh}@lsec.cc.ac.cn)

[‡] Department of Electrical Computer Science Engineering, University of Minnesota, Minneapolis, 55455, USA. (e-mail: luozq@umn.edu)

channel), analyze the complexity of the corresponding utility maximization problems, and propose a practical algorithm to solve them.

In addition to the aforementioned utility based formulations, various base station cooperation techniques have been proposed to mitigate inter-cell interferences, including multi-point coordinated transmission, or network multi-input multi-output (MIMO) transmission [7–16]. For example, the distributed or decentralized approaches are proposed for coordinated transmitter beamforming vectors in MISO interference channel in [7], [10], [13], [14], [16], [17], some of which are based on dual uplink channels. In particular, a distributed pricing algorithm for power control and beamformer design in the MISO interference networks is proposed in [17]. At each iteration, transmitters try to maximize their utility minus the total interference cost, i.e., the summation of the interference price times the received interference, and the interference price indicates the marginal decrease in the corresponding user's utility due to a marginal increase in interference associated with the particular receiver.

The references [11], [12] show that coordination enables the cellular network to enjoy a greater spectral efficiency. Most of these cooperative techniques require each base station to have not only full/partial CSI but also the knowledge of actual independent data streams to all remote terminals. With the complete sharing of data streams and CSI, the multi-cell scenario is effectively reduced to a single cell interference management problem with either total [18] or per-group-of-antenna power constraints [19], [20]. Among the major drawbacks of these techniques (in comparison to the utility based approaches) are their stringent requirement on base station coordination, the large demand on the communication bandwidth of backhaul links, as well as the heavy computational load associated with the increasing number of cells [21], [22]. The references [8], [15], [23], [24] provided characterizations of the achievable rate region and proved the existence of a unique Nash equilibrium which is inefficient in the sense that the achievable rates are bounded by a constant, regardless of the available transmit power. See [23], [25] for the recent results of the MISO channel.

Notation: The notation for this paper is as follows: Lower case boldface is used for column vectors. $(\cdot)^T$ and $(\cdot)^\dagger$ denote transpose and Hermitian transpose. (x, y) denotes a two-dimensional row vector. $\|\cdot\|$ denotes the Euclidean norm. Assuming $f(\mathbf{x})$ is a multi-variable function, $\nabla f(\mathbf{x})$ and $\nabla^2 f(\mathbf{x})$ denote its gradient and Hessian, respectively.

II. PROBLEM FORMULATION

Consider a cellular system in which there are K base stations each equipped with L transmit antennas. The K base stations wish to transmit respectively to K mobile receivers each having only a single antenna. Each base station can direct a beam to its intended receiver in such a way that the resulting interference to the other mobile units is small. Consider the single carrier case, and let $\mathbf{h}_{jk} \in \mathbb{C}^L$ denote L -dimensional complex channel vector between base station j and receiver k . Let $\mathbf{v}_k \in \mathbb{C}^L$ denote the beamforming vector used by

base station k , while s_k is a complex scalar denoting the information signal for user k with $\mathbb{E}|s_k|^2 = 1$. The transmitter vector of the j -th base station is $\mathbf{v}_j s_j$. Then the signal received by user k can be described as

$$y_k = \sum_{j=1}^K \mathbf{h}_{jk}^\dagger (\mathbf{v}_j s_j) + z_k, \quad 1 \leq k \leq K, \quad (1)$$

where z_k is the additive white Gaussian noise (AWGN) with variance $\sigma_k^2/2$ per real dimension. Treating interference as noise, we can write the SINR of each user as

$$\text{SINR}_k = \frac{|\mathbf{h}_{kk}^\dagger \mathbf{v}_k|^2}{\sigma_k^2 + \sum_{j \neq k} |\mathbf{h}_{jk}^\dagger \mathbf{v}_j|^2}. \quad (2)$$

Adopting an utility, we can formulate the optimal coordinated downlink beamforming problem as

$$\begin{aligned} \max \quad & H(r_1, r_2, \dots, r_K) \\ \text{s.t.} \quad & r_k = \log \left(1 + \frac{|\mathbf{h}_{kk}^\dagger \mathbf{v}_k|^2}{\sigma_k^2 + \sum_{j \neq k} |\mathbf{h}_{jk}^\dagger \mathbf{v}_j|^2} \right), \\ & \|\mathbf{v}_k\|^2 \leq P_k, \quad 1 \leq k \leq K, \end{aligned} \quad (3)$$

where P_k denotes the power budget of base station k , and $H(\cdot)$ denotes the system utility which may be any of the following

- Weighted sum-rate utility: $H_1 = \frac{1}{K} \sum_{k=1}^K w_k r_k$, with weight $w_k \geq 0$.
- Proportional fairness utility: $H_2 = \left(\prod_{k=1}^K r_k \right)^{1/K} \Leftrightarrow \frac{1}{K} \sum_{k=1}^K \log r_k$.
- Harmonic mean utility: $H_3 = K / \left(\sum_{k=1}^K r_k^{-1} \right)$.
- Min-rate utility: $H_4 = \min_{1 \leq k \leq K} r_k$.

According to [23], [24], problem (3) can be written in a more general form

$$\begin{aligned} \max \quad & H(r_1, r_2, \dots, r_K) \\ \text{s.t.} \quad & r_k = \log \left(1 + \frac{\mathbf{h}_{kk}^\dagger \mathbf{V}_k \mathbf{h}_{kk}}{\sigma_k^2 + \sum_{j \neq k} \mathbf{h}_{jk}^\dagger \mathbf{V}_j \mathbf{h}_{jk}} \right), \\ & \text{Trace}(\mathbf{V}_k) \leq P_k, \quad \mathbf{V}_k \succeq 0, \quad 1 \leq k \leq K, \end{aligned} \quad (4)$$

where \mathbf{V}_k is the transmit covariance matrix at transmitter k . The results in [23], [24] state that problem (4) has a rank-one optimal solution for each \mathbf{V}_k . This implies that problem (3) and problem (4) are equivalent. We focus on formulation (3) in this paper.

The above beamforming problem (3) can be nonconvex in general due to the nonlinear equality constraints. Given nonnegative weights (w_1, w_2, \dots, w_K) , the optimal tuple of rates (r_1, r_2, \dots, r_K) of (3) should lie on the boundary of the achievable rate region. See [8], [9], [15], [23], [24] for

various effort to characterize the achievable rate region of the interference channel.

In practice, the choice of utilities depends on a suitable compromise between system performance (total rates achievable) and user fairness. The sum-rate utility H_1 focuses entirely on system performance, while the min-rate utility H_4 places the highest emphasis on user fairness. The other two choices H_2 and H_3 represent an appropriate tradeoff between the two extremes.

III. COMPLEXITY ANALYSIS

In this section, we investigate the complexity status of the optimal coordinated downlink beamforming problem (3) under various choices of system utilities. We provide a complete analysis on when the problem is NP-hard and also identify subclasses of the problem that are solvable in polynomial time.

A. Computational complexity theory: a brief background

Generically, an optimization problem can be described by the minimization of an objective function over a feasible region. A decision version of the minimization problem is to decide if the feasible region contains a vector at which the objective function value is below a given threshold. The answer to the decision problem is binary, true or false, and there is no need to identify what the solution is. The decision version is typically easier to solve than the original optimization problem which requires the determination of an (globally) optimal solution. The size of an optimization problem instance is defined as the minimum length of a binary string required to describe the objective function and the feasible region. We say an algorithm solves the decision version of an optimization problem if for each instance of the problem, the algorithm correctly gives “true” or “false” answer. We can define the running time of an algorithm as the maximum number of basic computational steps (e.g., number of arithmetic operations) required to solve the decision version of an optimization problem of a given size. Typically, the algorithm’s running time is a function of the problem size.

In the computational complexity theory [26], [30], there are two important classes of optimization problems, P and NP. The class P contains optimization problems which are solvable (or decidable) by an algorithm whose running time grows at most as a polynomial function of the input size. It turns out that the class P is rather robust to the actual definition of input size. For example, P is invariant if we alternatively define the input size of an optimization problem as the sum of the problem dimension, the number of constraints and the binary length of the input data and threshold value. We say an algorithm is a pseudo polynomial time algorithm if its running time is a polynomial when we bound the size of numbers in the input by a constant. For example, many dynamic programming algorithms (e.g., Viterbi algorithm) are pseudo polynomial time since their state space are typically exponential if the numbers in the input are not bounded.

The class NP, which stands for *Nondeterministic Polynomial* time, consists of decision version of optimization problems whose “true” instances can be verified in polynomial time,

assuming the availability of a feasible solution that meets the threshold requirement. More formally, we say a nondeterministic algorithm solves a decision version of optimization problem if we can verify each “true” instance of the problem using a sequence of nondeterministic steps (i.e., involving random guesses). If the number of nondeterministic steps is polynomial, then the algorithm is said to have a nondeterministic polynomial running time. For example, for any symmetric matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ and a threshold value L , consider the problem of deciding if there exists a binary vector $\mathbf{x} \in \{-1, 1\}^n$ such that $\mathbf{x}^T \mathbf{Q} \mathbf{x} \leq L$. A nondeterministic algorithm to solve this problem is to guess a binary vector \mathbf{x} and then check if $\mathbf{x}^T \mathbf{Q} \mathbf{x} \leq L$ indeed holds. Such a binary vector \mathbf{x} exists for all “true” instances of the problem, and the verification process requires polynomially many steps although some steps may involve a random guess (of a component of \mathbf{x}). In this case, the problem is solvable in nondeterministic polynomial time. The class NP contains precisely those decision version of optimization problems that are solvable in nondeterministic polynomial time. Clearly, P is contained in NP. It is widely conjectured that $P \neq NP$, or equivalently, there are problems in NP which are not solvable in (deterministic) polynomial time.

A subset of problems in NP are called NP-complete. NP-complete problems are those in NP which are most difficult to solve, in the sense that if any one of them is in P, so is every other problem in NP. There are many well known NP-complete problems such as the traveling salesman problem, the 3-colorability problem, and so on. The latter problem is to decide if the nodes of a given graph can be colored in three colors so that each adjacent pair of nodes are colored differently. The 3-colorability problem is clearly in NP since we can easily check if a guessed coloring scheme meets the requirement. There is no known polynomial time algorithm to solve the 3-colorability problem. In fact, if this problem is solvable in polynomial time (i.e., in P), then every problem in NP is solvable in polynomial time, or equivalently $P=NP$. A problem \mathcal{P} is said to be NP-hard if it is at least as hard as those NP-complete problems, which means that the polynomial time solvability of \mathcal{P} would imply every NP-complete problem is in P. A NP-hard problem may not be in NP. For example, the binary quadratic minimization problem $\min_{\mathbf{x} \in \{-1, 1\}^n} \mathbf{x}^T \mathbf{Q} \mathbf{x}$ is NP-hard, since it is not known to be in NP, and is at least as hard as the NP-complete problem of deciding if there exists a binary vector \mathbf{x} such that $\mathbf{x}^T \mathbf{Q} \mathbf{x} \leq L$, where the threshold value L is given. A problem is called strongly NP-hard if it cannot be solved by a pseudo polynomial time algorithm unless $P=NP$.

To prove a problem \mathcal{P} is NP-complete, we need to show two things. First, we verify the problem is in NP. This step is usually easy. Second, we need to show \mathcal{P} is at least as difficult as a known NP-complete problem. This can be accomplished by a standard technique called polynomial time transformation. In a polynomial time transformation, we pick a known NP-complete problem and show that it is equivalent to a special case of \mathcal{P} . More precisely, we take an arbitrary instance of a known NP-complete problem, construct a special instance (with polynomial size) of \mathcal{P} , and then establish the equivalence of the two instances. To show a problem is NP-hard, we simply

ignore step 1, as there is no need to show \mathcal{P} is in NP. To show \mathcal{P} is strongly NP-hard, we need to additionally ensure the special instance of \mathcal{P} we construct involves only numbers of constant size, i.e., numbers whose combined bit length does not increase with the problem dimension or the number of constraints.

B. Maximization of the Weighted Sum-Rate Utility

Consider the system utility $H_1 = \frac{1}{K} \sum_{k=1}^K w_k r_k$. In the single antenna case ($L = 1$), the original system optimization problem (3) becomes the following (5), where $x_k = \|\mathbf{v}_k\|^2$, $\alpha_{jk} = \|\mathbf{h}_{jk}\|^2 / \|\mathbf{h}_{kk}\|^2$ and $\gamma_k = \sigma_k^2 / \|\mathbf{h}_{kk}\|^2$. Problem (5) is known to be NP-hard [1] even when the weights w_k are all equal, and the proof is based on a polynomial time transformation from the maximum independent set problem (which is known to be NP-complete). Thus, the general case of $L \geq 1$ is also NP-hard. For various special MISO channels, the sum-rate maximization problem can still be solved in polynomial time, see [25], [27–29].

C. Maximization of the Harmonic Mean Utility

We now study the complexity status of the optimal coordinated downlink beamforming problem (3) defined by the harmonic mean rate utility.

Theorem 3.1 (Harmonic Mean Utility): For the harmonic mean utility $H_3 = K / \left(\sum_{k=1}^K r_k^{-1} \right)$, the optimal coordinated downlink beamforming problem can be transformed into a convex optimization problem when $L = 1$, but is NP-hard when $L \geq 2$.

When there is only one transmit antenna ($L = 1$), the reference [1] shows that the harmonic mean rate maximization can be transformed into an equivalent convex problem. We thus focus on the case $L \geq 2$. Notice that the harmonic mean utility maximization problem is a continuous optimization problem. To show its NP-hardness, we need to transform a known NP-hard discrete problem to the harmonic mean maximization problem. To facilitate this transformation, it is necessary to induce certain discrete structure to the optimal solutions to the harmonic mean maximization problem. This is accomplished by using the concavity of the harmonic mean utility with respect to each beamforming vector \mathbf{v}_k . In particular, Lemma 3.2 (proved in Appendix I) shows that we can constrain the optimal beamforming vectors to be taken from two orthogonal vectors \mathbf{h}_a or \mathbf{h}_b .

The NP-hardness proof of Theorem 3.1 is based on a transformation from a variant of the 3-SAT [30] problem. To describe this variant, we need to define the UNANIMITY property and the NAE (stands for “not-all-equal”) property of a disjunctive clause¹.

Definition 3.1: For a given truth assignment to a set of Boolean variables, a disjunctive clause is said to be *UNANIMOUS* if all literals in the clause have the same value (whether

it is the *True* or the *False* value). Otherwise it is said to be satisfied in the *NAE* (Not-All-Equal) sense.

Notice that a disjunctive clause must be satisfied if it is to have the NAE property. We now define some *decision problems* over Boolean variables.

Definition 3.2: MAX-UNANIMITY problem: given a positive integer M and m disjunctive clauses defined over n Boolean variables, we ask whether there exists a truth assignment such that the number of *unanimous* disjunctive clauses is at least M . When the number of literals in each clause is two, we denote the corresponding problem as *MAX-2UNANIMITY problem*. When each clause contains three literals, the problem of determining whether there exists a truth assignment under which at least M clauses are satisfied in the NAE sense is called the *NAE-SAT problem*.

The NAE-SAT problem is known to be NP-complete [30]. Our next lemma says that the MAX-2UNANIMITY problem is also NP-complete.

Lemma 3.1: The MAX-2UNANIMITY problem is NP-complete.

Proof: We construct a polynomial time transformation from the NAE-SAT problem. It can be checked that the MAX-2UNANIMITY problem is in the class NP. Given a disjunctive clause with three literals, $c = x \vee y \vee z$, let us construct the following six clauses, each involving only two literals:

$$R(c) : x \vee \bar{y}, x \vee \bar{z}, y \vee \bar{x}, y \vee \bar{z}, z \vee \bar{x}, z \vee \bar{y}. \quad (6)$$

It can be checked that $R(c)$ has the following properties:

- 1) The number of unanimous clauses (i.e., all literals having the same value) in $R(c)$ is at most four.
- 2) The clause c is satisfied in the NAE sense if and only if the number of *unanimous* clauses in $R(c)$ is four.

Now given any instance ϕ of NAE-SAT problem, we construct a corresponding instance $R(\phi)$ of MAX-2UNANIMITY problem as follows: for each clause $c = \alpha \vee \beta \vee \gamma$ of ϕ , we add to $R(\phi)$ the six clauses in (6), with x, y, z replaced with the literals α, β, γ respectively. In this way, if ϕ has m clauses, then $R(\phi)$ will have $6m$ clauses. Let $M = 4m$. Then properties 1 and 2 imply that all clauses in ϕ are simultaneously satisfiable in the NAE sense if and only if at least $M = 4m$ clauses in $R(\phi)$ can be made *unanimous*. In particular, suppose that $4m$ clauses in $R(\phi)$ are unanimous under a given truth assignment. Since, by property 1, each group $R(c)$ of six clauses can have at most four unanimous clauses, it follows that exactly four clauses must be unanimous in each group. By property 2, this further implies that each clause in ϕ is satisfied in NAE sense. Conversely, any truth assignment that satisfies a clause c in the NAE sense will give rise to four unanimous clauses. Thus, if all m clauses in ϕ are satisfied in the NAE sense, then there will be $4m$ unanimous clauses in $R(\phi)$. Finally, this transformation is in polynomial time. ■

An immediate corollary of Lemma 3.1 is that the MAX-UNANIMITY problem is NP-complete. However, if we ask whether *all* of the m clauses (i.e., $M = m$) can be made unanimous, then the corresponding MAX-UNANIMITY problem, simply called UNANIMITY problem, can be solved in

¹Recall that for a given set of Boolean variables, a literal is defined as either a Boolean variable or its negation, while a disjunctive clause refers to a logical expression consisting of the logical “OR” of literals.

$$\max \sum_{k=1}^K w_k r_k, \quad \text{subject to } r_k = \log \left(1 + \frac{x_k}{\gamma_k + \sum_{j \neq k} \alpha_{jk} x_j} \right), \quad 0 \leq x_k \leq P_k, \quad 1 \leq k \leq K, \quad (5)$$

polynomial time (using a tree search technique). Also, notice that a disjunctive clause is *unanimous* under a given truth assignment if and only if c is not satisfied in *NAE* sense. This implies that MIN-3UNANIMITY problem is NP-complete. Since these results are not needed in the subsequent analysis, we state them without proof.

Proposition 3.1: The following is true:

- 1) MAX-UNANIMITY problem is NP-complete.
- 2) MIN-3UNANIMITY problem is NP-complete.
- 3) UNANIMITY problem is solvable in polynomial time.

We are now ready to prove Theorem 3.1. *Proof:* Let the utility in (3) be given by the harmonic mean function H_3 . Consider an instance of MAX-2UNANIMITY problem ϕ with clauses c_1, c_2, \dots, c_m defined over the variables x_1, x_2, \dots, x_n and an integer M . Let $\mathbf{h}_a = (1, 0)^T$, $\mathbf{h}_b = (0, 1)^T$ and $\mathbf{h} = (\sqrt{N}, 0)^T$, where N is a large positive number (to be specified later). We write each clause $c_j = \alpha_j \vee \beta_j$, with α_j, β_j taken from $\{x_1, x_2, \dots, x_n, \bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$. Let us define two mappings

$$\pi, \tau : \{1, 2, \dots, m\} \mapsto \{\pm 1, \pm 2, \dots, \pm n\}$$

such that

$$\pi(j) = \begin{cases} i, & \text{if } \alpha_j = x_i, \\ -i, & \text{if } \alpha_j = \bar{x}_i, \end{cases} \quad \text{and } \tau(j) = \begin{cases} i, & \text{if } \beta_j = x_i, \\ -i, & \text{if } \beta_j = \bar{x}_i. \end{cases}$$

For instance, if $c_4 = x_3 \vee \bar{x}_5$, then we have $\alpha_4 = x_3, \beta_4 = \bar{x}_5$, with $\pi(4) = 3$ and $\tau(4) = -5$. For $i = \pm 1, \pm 2, \dots, \pm n$, we define

$$\mathbf{h}_i = \begin{cases} \mathbf{h}_a, & \text{if } i > 0, \\ \mathbf{h}_b, & \text{if } i < 0. \end{cases}$$

Given an instance of MAX-2UNANIMITY problem, we construct the following (7) as an instance of (3) (the inverse of harmonic mean utility minimization is equivalent to the harmonic mean rate utility maximization) with a total of $K = 4n + 2m$ users. Herein, each Boolean variable x_i corresponds to four users, including user $4i$ (called “variable user”) and user $4i - 1, 4i - 2$ and $4i - 3$ (called “auxiliary variable users”); while each clause c_j corresponds to a pair of users, i.e., user $4n + 2j$ and $4n + 2j - 1$ (called “clause users”). In our construction (7), each (variable, auxiliary variable or clause) user k is associated with a transmitter beamforming vector $\mathbf{v}_k, k = 1, 2, \dots, 4n + 2m$.

In (7), the n “variable users” communicate interference free. Their channel vectors are $(\sqrt{0.9}, \sqrt{0.9})^T$ and their noise power are 1. The $3n$ “auxiliary variable users” $4i - 1, 4i - 2, 4i - 3, i = 1, 2, \dots, n$, do suffer from crosstalk interference from the “variable user” $4i$. That is, the interference channel vectors from the “variable user” $4i$ are $(1, 1)^T, (1, 0)^T, (0, 1)^T$, respectively; the direct link channel vectors are $(1, 0)^T, (10, 0)^T$ and $(10, 0)^T$; the self noise power at each “auxiliary variable user” is zero. For the “clause users” $4n + 2j$ and $4n + 2j - 1, j = 1, 2, \dots, m$, their

channel vectors are $\mathbf{h} = (\sqrt{N}, 0)^T$ and their noise powers are 1. Let $\mathbf{v}_{4i}, i = 1, 2, \dots, n$, denote the transmit beamforming vector of the “variable user” $4i$. For example, the clause $c_1 = x_2 \vee \bar{x}_3$ corresponds to $j = 1$ and is associated with two “clause users” which are denoted by $4n + 1, 4n + 2$. Since $\pi(1) = 2$ and $\tau(1) = -3$, the two users $4n + 1$ and $4n + 2$ experience interferences from “variable users” $4|\pi(1)| = 8$ and $4|\tau(1)| = 12$. The corresponding interference terms for these two “clause users” $4n + 1, 4n + 2$ are $|\mathbf{h}_a^\dagger \mathbf{v}_8|^2 + |\mathbf{h}_b^\dagger \mathbf{v}_{12}|^2$ and $|\mathbf{h}_b^\dagger \mathbf{v}_8|^2 + |\mathbf{h}_a^\dagger \mathbf{v}_{12}|^2$, respectively.

The correspondence between MAX-2UNANIMITY problem and the optimal coordinated downlink beamforming problem (7) is listed in Table I. Notice that r_{4n+2j} can be obtained from clause c_j according to Table I and $r_{4n+2j-1}$ can be obtained from r_{4n+2j} by swapping \mathbf{h}_a with \mathbf{h}_b .

TABLE I
VARIABLE CORRESPONDENCE

MAX-2UNANIMITY	Problem (7)
variable x_i	beamforming vector \mathbf{v}_{4i}
clause c_j	rates r_{4n+2j} and $r_{4n+2j-1}$
literal x_i	interference term $ \mathbf{h}_a^\dagger \mathbf{v}_{4i} ^2$
literal \bar{x}_i	interference term $ \mathbf{h}_b^\dagger \mathbf{v}_{4i} ^2$

We first fix some easy variables of (7) to simplify the problem. Since each of the beamforming vectors $\mathbf{v}_{4n+2j}, \mathbf{v}_{4n+2j-1}, j = 1, 2, \dots, m$, and $\mathbf{v}_{4i-l}, i = 1, 2, \dots, n, l = 1, 2, 3$, appears exactly once in the objective function, it follows that, by optimality, these beamforming vectors must be matched to the corresponding channel vectors. That is, $\mathbf{v}_{4n+2j}^* = \mathbf{v}_{4n+2j-1}^* = (1, 0)^T$, and $\mathbf{v}_{4i-1}^* = \mathbf{v}_{4i-2}^* = \mathbf{v}_{4i-3}^* = (1, 0)^T$. Substituting these optimal beamforming vectors into (7) yields (8). It only remains to determine the optimal beamforming vectors $\mathbf{v}_{4i}^*, i = 1, 2, \dots, n$. For this purpose, we need the following key lemma whose proof is relegated to Appendix I.

Lemma 3.2: If $N \geq 2(e^{40m} - 1)$, then the optimal beamforming vectors $\{\mathbf{v}_{4i}^*\}$ for the optimization problem (8) must be either $\mathbf{h}_a = (1, 0)^T$ or $\mathbf{h}_b = (0, 1)^T$.

It follows from Lemma 3.2 that the optimal beamforming vector \mathbf{v}_{4i}^* of (8) is either \mathbf{h}_a or \mathbf{h}_b . In either case, the sum of inverse rates

$$\frac{1}{r_{4i}} + \frac{1}{r_{4i-1}} + \frac{1}{r_{4i-2}} + \frac{1}{r_{4i-3}} = \frac{1}{\log 1.9} + \frac{1}{\log 101} + \frac{1}{\log 2} \triangleq C$$

is a constant C . Thus, regardless of whether $\mathbf{v}_{4i}^* = \mathbf{h}_a$ or \mathbf{h}_b , the first sum in the objective function of (8) remains unchanged and equals nC . Thus, we only need to consider the second sum in the objective function of (8). Notice that

$$\begin{aligned}
\min \quad & \sum_{i=1}^n \left(\frac{1}{r_{4i}} + \frac{1}{r_{4i-1}} + \frac{1}{r_{4i-2}} + \frac{1}{r_{4i-3}} \right) + \sum_{j=1}^m \left(\frac{1}{r_{4n+2j}} + \frac{1}{r_{4n+2j-1}} \right) \\
\text{s.t.} \quad & r_{4i} = \log \left(1 + |(\sqrt{0.9}, \sqrt{0.9}) \mathbf{v}_{4i}|^2 \right), \\
& r_{4i-1} = \log \left(1 + \frac{|(1, 0) \mathbf{v}_{4i-1}|^2}{|(1, 1) \mathbf{v}_{4i}|^2} \right), \\
& r_{4i-2} = \log \left(1 + \frac{|(10, 0) \mathbf{v}_{4i-2}|^2}{|(1, 0) \mathbf{v}_{4i}|^2} \right), \\
& r_{4i-3} = \log \left(1 + \frac{|(10, 0) \mathbf{v}_{4i-3}|^2}{|(0, 1) \mathbf{v}_{4i}|^2} \right), \\
& r_{4n+2j} = \log \left(1 + \frac{|\mathbf{h}^\dagger \mathbf{v}_{4n+2j}|^2}{1 + |\mathbf{h}_{\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2} \right), \\
& r_{4n+2j-1} = \log \left(1 + \frac{|\mathbf{h}^\dagger \mathbf{v}_{4n+2j-1}|^2}{1 + |\mathbf{h}_{-\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{-\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2} \right), \\
& \|\mathbf{v}_k\|^2 \leq 1, \quad 1 \leq i \leq n, \quad 1 \leq j \leq m, \quad 1 \leq k \leq 4n + 2m.
\end{aligned} \tag{7}$$

$$\begin{aligned}
\min \quad & \sum_{i=1}^n \left(\frac{1}{r_{4i}} + \frac{1}{r_{4i-1}} + \frac{1}{r_{4i-2}} + \frac{1}{r_{4i-3}} \right) + \sum_{j=1}^m \left(\frac{1}{r_{4n+2j}} + \frac{1}{r_{4n+2j-1}} \right) \\
\text{s.t.} \quad & r_{4i} = \log \left(1 + 0.9|(1, 1) \mathbf{v}_{4i}|^2 \right), \\
& r_{4i-1} = \log \left(1 + 1/|(1, 1) \mathbf{v}_{4i}|^2 \right), \\
& r_{4i-2} = \log \left(1 + 100/|(1, 0) \mathbf{v}_{4i}|^2 \right), \\
& r_{4i-3} = \log \left(1 + 100/|(0, 1) \mathbf{v}_{4i}|^2 \right), \\
& r_{4n+2j} = \log \left(1 + N / \left(1 + |\mathbf{h}_{\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2 \right) \right), \\
& r_{4n+2j-1} = \log \left(1 + N / \left(1 + |\mathbf{h}_{-\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{-\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2 \right) \right), \\
& \|\mathbf{v}_{4i}\|^2 \leq 1, \quad 1 \leq i \leq n, \quad 1 \leq j \leq m.
\end{aligned} \tag{8}$$

$$\frac{1}{r_{4n+2j}} + \frac{1}{r_{4n+2j-1}} = \begin{cases} \frac{1}{\log(1 + N/3)} + \frac{1}{\log(1 + N/1)}, & \text{if } c_j \text{ is unanimous,} \\ \frac{1}{\log(1 + N/2)} + \frac{1}{\log(1 + N/2)}, & \text{if } c_j \text{ is not unanimous.} \end{cases}$$

the value of each term in the second sum only depends on whether clause c_j is unanimous (see the second equation in the next page). Since

$$\frac{1}{\log(1 + N/3)} + \frac{1}{\log(1 + N)} < \frac{2}{\log(1 + N/2)}$$

from Claim 1 in Appendix I, it follows that second sum of (8) will be smaller if more clauses are satisfied unanimously. Therefore, the minimum of (8) is only related to the maximum number of unanimous clauses in the given MAX-2UNANIMITY problem. Specifically, the minimum of (8) is no more than

$$nC + \frac{M}{\log(1 + N/3)} + \frac{M}{\log(1 + N)} + \frac{2(m - M)}{\log(1 + N/2)} \tag{9}$$

if and only if there exists an appropriate truth assignment such that at least M clauses are made unanimous for the given MAX-2UNANIMITY problem. Thus, we have transformed the problem of MAX-2UNANIMITY problem to the problem of checking if problem (7) will have an optimal value below the above threshold (9).

Finally, given an instance of the MAX-2UNANIMITY problem, we can construct the harmonic mean rate maximization problem (7) in polynomial time. Since the MAX-2UNANIMITY problem is NP-complete (Lemma 3.1), it follows that the optimal coordinated beamforming problem (3) with harmonic mean utility is NP-hard. ■

A few remarks are in order. First, it follows from the proof of Theorem 3.1 that even if the optimal transmit power levels are known (i.e., $\|\mathbf{v}_k\|^2 \leq P_k$ is replaced with $\|\mathbf{v}_k\|^2 = P_k$), the problem of finding the optimal beamforming directions of harmonic mean rate maximization problem is still NP-hard. Second, we have set the noise powers of users $4i - 1, 4i - 2, 4i - 3, i = 1, 2, \dots, n$, to zero in (7). These settings simplify the proof and do not reduce any generality. We could have used small noise power values in the proof (even though some extra argument is needed), since there is a positive gap between the global optimal value and the local optimal values of (7). Finally, our proof actually implies that there is a *positive* probability (measure) that a randomly generated MISO coordinated beamforming problem under the harmonic

mean utility is NP-hard to solve. In particular, by continuity, all slightly perturbed versions of the constructed instance (7) (i.e., channel vectors, noise/transmit powers are slightly changed) will be equivalent to the MAX-2UNANIMITY problem. This is because there is a positive (and constant) jump in the global optimal value of the constructed example when the optimal value of the corresponding MAX-2UNANIMITY problem increases by one. When channel conditions change slightly, this one-to-one correspondence between the optimal values of the two problems and the property of the discrete jump in the optimal value of the constructed MISO problem remain valid.

D. Maximization of Proportional Fairness Utility

Like the harmonic mean utility, we have the following hardness result.

Theorem 3.2 (Proportional Fairness Utility): For the proportional fairness utility $H_2 = \left(\prod_{k=1}^K r_k \right)^{1/K}$, the optimal coordinated downlink beamforming problem can be transformed into a convex optimization problem when $L = 1$, but is NP-hard when $L \geq 2$.

Proof: The first part of Theorem 3.2 is proved in [1]. For the second part, the argument is similar to that of Theorem 3.1. We only give a proof outline below.

First, we have the following lemma whose proof is provided in Appendix II.

Lemma 3.3: The function $f(x) = \log \log \left(1 + \frac{1}{\sigma^2 + x} \right)$ is strictly convex in $x \geq 0$ for any σ .

Second, given any MAX-2UNANIMITY problem, an instance (10) of the optimal coordinated downlink beamforming problem (3) with utility $\log H_2$ (equivalent to proportional fairness utility maximization) and $3n + 2m$ users is constructed as follows. Notice that each global optimal solution of (10) must have $\mathbf{v}_{3i-1}^* = \mathbf{v}_{3i-2}^* = (1, 0)^T$, $i = 1, 2, \dots, n$, and $\mathbf{v}_{3n+2j}^* = \mathbf{v}_{3n+2j-1}^* = (1, 0)^T$, $j = 1, 2, \dots, m$. Moreover, we consider the following parametric optimization problem (similar to (19) in the harmonic mean case):

$$\begin{aligned} \max \quad & \log r_3 + \log r_2 + \log r_1 \\ \text{s.t.} \quad & r_3 = \log \left(1 + |(\sqrt{0.1}, \sqrt{0.1})\mathbf{v}_3|^2 \right), \\ & r_2 = \log \left(1 + 1/(\sigma^2 + |(1, 0)\mathbf{v}_3|^2) \right), \\ & r_1 = \log \left(1 + 1/(\sigma^2 + |(0, 1)\mathbf{v}_3|^2) \right), \\ & \|\mathbf{v}_3\| = t, \end{aligned} \quad (11)$$

where $\sigma > 0$ is a constant and t is a parameter. The global maxima of (11) should be $(t, 0)^T$ and $(0, t)^T$ when σ is small. Furthermore, the optimum value of (11) is an increasing function with respect to $t \in [0, 1]$. Using an argument similar to that of the harmonic mean case for (8), each globally optimal beamforming solution \mathbf{v}_{3i}^* of (10) should be either \mathbf{h}_a or \mathbf{h}_b when $N \geq 3(e^{6m} - 1)$. When restricted to solutions of the form $\mathbf{v}_{3i}^* = \mathbf{h}_a$ or \mathbf{h}_b , the maximum of (10) is only linearly related to the maximum number of unanimous clauses in the given MAX-2UNANIMITY problem. Thus, maximizing the number of unanimous clauses is the same as solving (10). Since MAX-2UNANIMITY problem is NP-complete

(Lemma 3.1), it follows that the optimal coordinated downlink beamforming problem with utility H_2 is also NP-hard. ■

E. Maximization of Min-Rate Utility

Let the system utility function be given by $H = H_4$. In this case, the problem can be solved in polynomial time for arbitrary L and K [1], [2]. Specifically, letting

$$r = \min_{1 \leq k \leq K} \{r_k\},$$

the min-rate utility maximization problem becomes

$$\begin{aligned} \max \quad & r \\ \text{s.t.} \quad & r \leq \log \left(1 + \frac{|\mathbf{h}_{kk}^\dagger \mathbf{v}_k|^2}{\sigma_k^2 + \sum_{j \neq k} |\mathbf{h}_{jk}^\dagger \mathbf{v}_j|^2} \right), \\ & \|\mathbf{v}_k\|^2 \leq P_k, \quad 1 \leq k \leq K. \end{aligned} \quad (12)$$

Given a $r \geq 0$, we can efficiently check if there exists $\mathbf{v}_k, k = 1, 2, \dots, K$, such that the constraints in (12) are satisfied. This feasibility problem is a second order cone programming, which can be solved efficiently using interior-point methods. The following theorem is a generalization of the result of [2], which deals with the single-cell case.

Theorem 3.3 (Min-Rate Utility): For the min-rate utility, the optimal coordinated downlink beamforming problem can be solved in polynomial time with arbitrary K and L .

Proof: We give the following bisection algorithm for solving (12).

A Polynomial Time Algorithm for Min-Rate Utility Maximization

- Step 1.** Initialization: Choose r_ℓ and r_u such that the optimal r_{opt} lies in $[r_\ell, r_u]$ and a tolerance ϵ .
Step 2. If $r_u - r_\ell \leq \epsilon$, stop, else go to **Step 3**.
Step 3. Let $r_{mid} = (r_\ell + r_u)/2$ and solve an SOCP problem to check the feasibility problem of (12) with $r = r_{mid}$. If feasible, let $r_\ell = r_{mid}$, else set $r_u = r_{mid}$ and go to **Step 2**.

According to standard analysis of path-following interior-point methods, **Step 3** can be finished in $O(K^{3.5}L^{3.5})$ time. With regards to initial choices for r_ℓ and r_u , we can let $\bar{\mathbf{v}}_k = \mathbf{h}_{kk}\sqrt{P_k}/\|\mathbf{h}_{kk}\|$ (space matched beamformer) and

$$\begin{aligned} r_\ell &= \min_k \log \left(1 + \frac{|\mathbf{h}_{kk}^\dagger \bar{\mathbf{v}}_k|^2}{\sigma_k^2 + \sum_{j \neq k} |\mathbf{h}_{jk}^\dagger \bar{\mathbf{v}}_j|^2} \right), \\ r_u &= \min_k \log \left(1 + \frac{|\mathbf{h}_{kk}^\dagger \bar{\mathbf{v}}_k|^2}{\sigma_k^2} \right). \end{aligned}$$

It takes $\log_2((r_u - r_\ell)/\epsilon)$ iterations to reach tolerance ϵ . Thus, a total of $O(K^{3.5}L^{3.5} \log_2((r_u - r_\ell)/\epsilon))$ arithmetic operations are needed in the worst case. Since

$$r_u - r_\ell \leq \log(1 + P_k \|\mathbf{h}_{kk}\|^2 / \sigma_k^2) \leq P_k \|\mathbf{h}_{kk}\|^2 / \sigma_k^2,$$

the second equation in the next page holds true, which is a polynomial in the length of input data (all the channel vectors \mathbf{h}_{kj} , noise power σ_k^2 and power upper bound P_k). Thus, the

$$\begin{aligned}
& \max \quad \sum_{i=1}^n (\log r_{3i} + \log r_{3i-1} + \log r_{3i-2}) + \sum_{j=1}^m (\log r_{3n+2j} + \log r_{3n+2j-1}) \\
& \text{s.t.} \quad r_{3i} = \log \left(1 + |(\sqrt{0.1}, \sqrt{0.1}) \mathbf{v}_{3i}|^2 \right), \\
& \quad r_{3i-1} = \log \left(1 + \frac{|(1, 0) \mathbf{v}_{3i-1}|^2}{|(1, 0) \mathbf{v}_{3i}|^2} \right), \\
& \quad r_{3i-2} = \log \left(1 + \frac{|(1, 0) \mathbf{v}_{3i-2}|^2}{|(0, 1) \mathbf{v}_{3i}|^2} \right), \\
& \quad r_{3n+2j} = \log \left(1 + \frac{|\mathbf{h}^\dagger \mathbf{v}_{3n+2j}|^2}{1 + |\mathbf{h}_{\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2} \right), \\
& \quad r_{3n+2j-1} = \log \left(1 + \frac{|\mathbf{h}^\dagger \mathbf{v}_{3n+2j-1}|^2}{1 + |\mathbf{h}_{-\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{-\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2} \right), \\
& \quad \|\mathbf{v}_k\|^2 \leq 1, \quad 1 \leq i \leq n, \quad 1 \leq j \leq m, \quad 1 \leq k \leq 3n + 2m.
\end{aligned} \tag{10}$$

$$K^{3.5} L^{3.5} \log_2 ((r_u - r_\ell) / \epsilon) \leq K^{3.5} L^{3.5} (|\log_2 P_k| + 2|\log_2 \|\mathbf{h}_{kk}\|| + 2|\log_2 \sigma_k| + |\log_2 \epsilon|),$$

above algorithm has a polynomial time worst case complexity. ■

Remark: The algorithm described herein easily extend to the weighted min-rate maximization problems. In [28], the weighted min-rate problem is related to the weighted sum MSE minimization problem and the weighted sum-rate maximization problems through the nonnegative matrix theory.

Table II summarizes the complexity status of the optimal coordinated downlink beamforming problem (3) for different choices of utilities.

IV. A CYCLIC DESCENT ALGORITHM

In this section, we consider numerical algorithms for the coordinated beamforming problem (3) with harmonic mean utility H_3 :

$$\begin{aligned}
& \min \quad \rho(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K) \\
& \text{s.t.} \quad \|\mathbf{v}_k\|^2 \leq P_k, \quad 1 \leq k \leq K,
\end{aligned} \tag{13}$$

where

$$\begin{aligned}
\rho(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K) &= \sum_{k=1}^K \frac{1}{r_k}, \\
r_k &= \log \left(1 + \frac{|\mathbf{h}_{kk}^\dagger \mathbf{v}_k|^2}{\sigma_k^2 + \sum_{j \neq k} |\mathbf{h}_{jk}^\dagger \mathbf{v}_j|^2} \right).
\end{aligned}$$

Since this problem is NP-hard (proved in Section III), we are led to develop efficient algorithms to find a high quality approximate solution or a stationary point for (13). Due to variable *separability* in the constraints of (13) and our desire for distributed implementation, we propose to solve (13) by *cyclicly* adjusting the beamforming vector \mathbf{v}_k while assuming the beamforming vectors $\{\mathbf{v}_j : j \neq k\}$ are fixed. In other words, we solve a sequence of per-base station problems

$$\begin{aligned}
& \min_{\mathbf{v}_k} \quad \rho(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K) \\
& \text{s.t.} \quad \|\mathbf{v}_k\|^2 \leq P_k.
\end{aligned} \tag{14}$$

Interestingly, the pricing algorithm introduced in [17] can be viewed as a partially linearized version of our cyclic coordinate descent algorithm. In particular, our proposed algorithm tries to allocate resources of the k -th transmitter by maximizing the summation of all users' utility functions subject to its power constraint ($\|\mathbf{v}_k\|^2 \leq P_k$); while the pricing algorithm lets transmitter k maximize its own utility function plus the summation of the first order approximation of all other users' utility functions at the current point $I_{kj} = |\mathbf{h}_{kj}^\dagger \mathbf{v}_k|^2$, where I_{kj} denotes the interference at the j -th receiver from the k -th transmitter.

A. An Inexact Cyclic Coordinate Descent Algorithm

The cyclic coordinate descent algorithm is also known as the nonlinear Gauss-Seidel iteration [32]. There are several studies of this type of algorithms [32–36] with many applications in engineering [37]. However, most of these studies require either the convexity of objective function or exact solution of subproblems (14), which not only is costly but also may result in algorithm divergence [33]. Below we consider a general differentiable optimization problem with separable constraints

$$\begin{aligned}
& \min \quad f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K) \\
& \text{s.t.} \quad \mathbf{x}_k \in X_k, \quad 1 \leq k \leq K,
\end{aligned} \tag{15}$$

where the feasible set $X := \prod_{k=1}^K X_k$ is separable, bounded and closed. We propose an easily implementable cyclic coordinate descent algorithm which simply requires a sufficient decrease in the objective of (14) at each iteration. The algorithm can be applied to solve the utility maximization problem (3) with $H = H_1, H_2$ and H_3 and have the same convergence properties because they have smooth objective functions and a separable feasible region. But the same can not be said about H_4 since it is non-differentiable.

TABLE II
COMPLEXITY STATUS OF THE OPTIMAL COORDINATED DOWNLINK BEAMFORMING PROBLEM

Utility Class	Weighted Sum-Rate	Proportional Fairness	Harmonic Mean	Min-Rate
$L = 1$, any K	NP-hard [1]	Convex [1]	Convex [1]	Poly. time Algorithm [1], [2]
$L \geq 2$, any K	NP-hard	NP-hard	NP-hard	Poly. time Algorithm [1], [2]

An Inexact Cyclic Coordinate Descent Algorithm

Step 1. Initialization: choose $\mathbf{x}^1 = [\mathbf{x}_1^1, \mathbf{x}_2^1, \dots, \mathbf{x}_K^1]$ and a tolerance $\epsilon > 0$.

Step 2. Iteration $i \geq 1$: Denote for $k = 1, 2, \dots, K$,

$$\mathbf{z}_k^{i+1} = (\mathbf{x}_1^{i+1}, \dots, \mathbf{x}_{k-1}^{i+1}, \mathbf{x}_k^{i+1}, \mathbf{x}_{k+1}^i, \dots, \mathbf{x}_K^i), \quad (16)$$

and let $\mathbf{z}_0^{i+1} = \mathbf{x}^i$.

For $k = 1, 2, \dots, K$,

- Compute the gradient projection direction for the component \mathbf{x}_k according to

$$\mathbf{d}_k^{i+1} = P_{X_k}[\mathbf{x}_k^i - \nabla_{\mathbf{x}_k} f(\mathbf{z}_{k-1}^{i+1})] - \mathbf{x}_k^i, \quad (17)$$

where $P_{X_k}[\cdot]$ denotes the orthogonal projection to X_k .

- Determine a stepsize α_k^{i+1} using the backtracking line search technique [38].
- Update $\mathbf{x}_k^{i+1} = \mathbf{x}_k^i + \alpha_k^{i+1} \mathbf{d}_k^{i+1}$.

Let $\mathbf{x}^{i+1} = \mathbf{z}_K^{i+1}$.

Step 3. Termination: If $\|\mathbf{x}^{i+1} - \mathbf{x}^i\| \leq \epsilon$, then stop. Else, set $i = i + 1$ go to **Step 2**.

When specialized to the MISO downlink beamforming problem, the above inexact cyclic optimization procedure can be implemented in a distributed fashion. At the initial step, each base station needs to know the local CSI for all channels originating from that transmitter (either through feedback or reverse-link estimation [31]). The only information to be exchanged are the SINR terms at K receivers. In subsequent iterations, a base station updates its beamforming vector by solving (14) inexactly using a gradient projection algorithm. After that, all receivers measure individual SINR terms and send the SINR information to the next base station. The inexact cyclic coordinate descent algorithm enables each transmitter to update its beamformer with only limited information exchange.

The next result shows that the above Inexact Cyclic Coordinate Descent Algorithm converges to a KKT point of (15). The proof of this result is relegated to Appendix III.

Theorem 4.1: Suppose $f(\mathbf{x})$ is twice continuously differentiable and bounded below, and the feasible set $X := \prod_{k=1}^K X_k$ is convex, separable and compact. Then every accumulation point of the sequence $\{\mathbf{x}^i\}$ generated by the inexact cyclic coordinate descent algorithm is a stationary point of (15).

The separability of constraints is necessary for the algorithm's convergence. The following example (taken from [35]) shows that, without the separability, the algorithm can get

stuck at an uninteresting point:

$$\begin{aligned} \min \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 2. \end{aligned}$$

This strongly convex function has a unique global solution at $x_1^* = x_2^* = 1$. However, if the initial point is (1.5, 0.5), the cyclic coordinate descent algorithm will be stuck.

Specializing the inexact cyclic coordinate descent algorithm to the coordinated beamforming problem (13), we need to perform a projected gradient descent iteration for the subproblem (14). In this case, we have a ball constraint $V_k = \{\mathbf{v} \mid \|\mathbf{v}\|^2 \leq P_k\}$ and the corresponding projection is straightforward

$$P_{V_k}(\mathbf{v}) = \begin{cases} \mathbf{v}, & \text{if } \|\mathbf{v}\|^2 \leq P_k, \\ \frac{\sqrt{P_k} \mathbf{v}}{\|\mathbf{v}\|}, & \text{if } \|\mathbf{v}\|^2 > P_k. \end{cases} \quad (18)$$

As a variant, we can also use the so called Barzilai-Borwein (BB) projection step for the subproblem (14) to replace the standard gradient projection step. In particular, at iteration i , the BB gradient projection direction \mathbf{d}_{BB}^i is given by

$$\begin{cases} \mathbf{d}_{BB}^i &= P_{V_k}(\mathbf{v}_k^i - \alpha_{BB}^i \nabla_{\mathbf{v}_k} \rho(\mathbf{v}^i)) - \mathbf{v}_k^i, \\ \alpha_{BB}^i &= \frac{\|\mathbf{s}^{i-1}\|^2}{(\mathbf{s}^{i-1})^T \mathbf{y}^{i-1}}, \end{cases}$$

where $\mathbf{s}^{i-1} = \mathbf{v}_k^i - \mathbf{v}_k^{i-1}$, $\mathbf{y}^{i-1} = \nabla_{\mathbf{v}_k} \rho(\mathbf{v}^i) - \nabla_{\mathbf{v}_k} \rho(\mathbf{v}^{i-1})$. It can be shown that the above BB gradient projection direction is always a descent direction. The R -linear convergence of the BB method has been established for strongly convex quadratic functions in [39].

If we use a single gradient projection step to inexactly solve the partially linearized subproblems in the pricing algorithm [17], then the resulting inexact pricing algorithm is algorithmically identical to the inexact cyclic coordinate descent algorithm considered herein. This is because the gradient vectors of the two utility functions in the respective subproblems (i.e., $\rho(\mathbf{v})$ and its partially linearized versions) are exactly the same. This observation, coupled with Theorem 4.1, immediately implies the convergence (to a KKT solution) of the inexact pricing algorithm for MISO interference channel. The latter is interesting since the convergence of the original pricing algorithm has not yet been established.

V. NUMERICAL SIMULATIONS

To evaluate the effectiveness of the cyclic descent algorithm, we have conducted numerical simulations for a 7-cell network with one user per cell as shown in Fig. 1. Each base station is equipped with L antennas. Similar to [5], standard WiMax parameters are used in all the simulations; see Table III, where d is the distance in kilometers. The location of each remote

TABLE III
STANDARD WiMAX PARAMETERS

Model or Parameters	Values
noise power spectral density	-162 dBm/Hz
path loss model	$128.1 + 37.6 \log_{10}(d)$
log-normal shadowing	8 dB
distance between neighboring base stations	2.8km
antenna gain	15 dBi

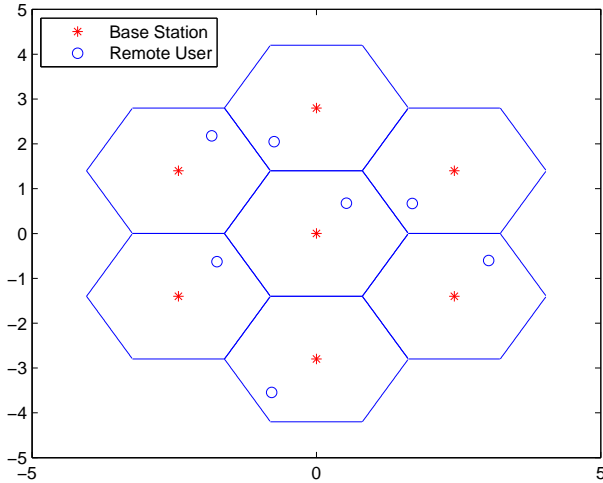


Fig. 1. A wireless network with seven base stations and single user per cell.

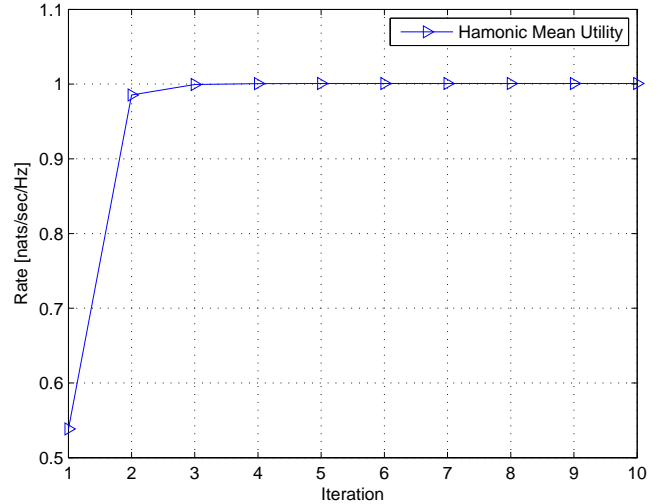


Fig. 2. A typical iteration process of the cyclic coordinate descent algorithm with $K = 7$, $L = 4$ and $P = 30$ dBm.

user is chosen randomly within its cell such that it is at least 0.5km away from the corresponding base station.

Figure 2 plots the iteration process of BB projection method for the coordinated downlink beamforming problem with harmonic mean utility. It can be seen that most of improvement is achieved in the first 1-2 iterations, making the method attractive for practical implementations.

For a two-user MISO channel, a parametrization of the achievable rate region boundary was given in [8], [24]. We can use this parametrization to compute the global optimal solution of the coordinated MISO downlink beamforming problem by searching along the rate region boundary. In Fig. 3, the performance of our proposed cyclic descent algorithm is compared against the global optimum for 50 randomly generated two-user MISO channel. It can be seen that the proposed algorithm either achieves, or nearly achieves, the global optimality.

Figure 4 shows the performance achieved by the cyclic coordinate descent algorithm versus different number of antennas for various utility functions and with a fixed transmit power $P = 30$ dBm at each base station. Each point in Fig. 4 is obtained by averaging over 500 independent channel realizations. Space matched beamformers are used as the benchmark. It can be seen that the transmit rates improve significantly over the benchmark solution. When the number of users ($K = 7$) and the transmit power ($P = 30$ dBm) are fixed, the utility increases linearly with the number of antennas L , suggesting an additive system gain with increasing L .

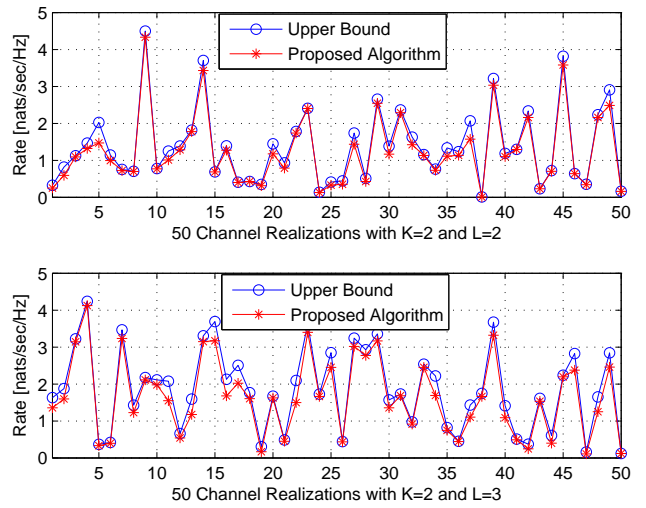


Fig. 3. Performance comparison of the proposed algorithm and the parametrization method with $K = 2$, $P = 30$ dBm.

VI. CONCLUSION

Coordinated transmit beamforming is a promising approach for interference mitigation in a MISO interference channel. A major design challenge is to find, for a given channel state, a globally optimal beamforming strategy under an appropriate utility criterion. In the single carrier case with a single antenna

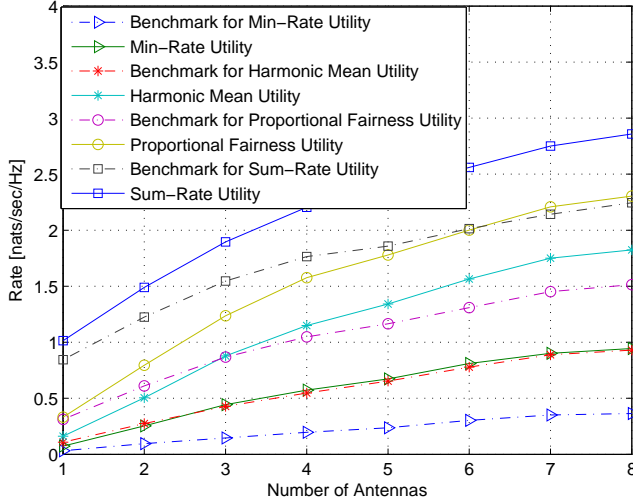


Fig. 4. Performance comparison of coordinated beamformers with space matched beamformers with $K = 7$, $P = 30$ dBm.

per transmitter, maximizing the (weighted) sum-rate is known to be NP-hard. However, the same problem is polynomial time solvable when the proportional fairness, harmonic mean or max-min utility is used. It turns out the situation with multiple transmit antennas is rather different. For instance, when each transmitter (e.g., a base station) is equipped with two antennas, the corresponding joint beamformer design problem becomes NP-hard under either the proportional fairness or the harmonic mean criterion. These complexity results suggest that we should abandon effort to find globally optimal beamformers for a general MISO interference channel unless the max-min utility is used. In the latter case the problem remains solvable in polynomial time.

Motivated by these complexity results, we propose a simple distributed inexact cyclic coordinate descent algorithm to find a (locally optimal) beamforming strategy. Our algorithm exploits the separable structure of the power constraints, and is provably globally convergent to a KKT solution. This algorithm requires only local CSI and an exchange of local SINRs at each iteration. Numerical experiments with WiMax system parameters show that the proposed algorithm is both effective and efficient, providing significant rate gain over the space matched beamforming strategy.

VII. ACKNOWLEDGMENTS

The authors wish to thank Professor Wei Yu of University of Toronto for his help in numerical simulations.

APPENDIX I PROOF OF LEMMA 3.2

The proof consists of establishing three claims.

Claim 1: The function $\log^{-1} \left(1 + (\sigma^2 + x)^{-1} \right)$ is strictly concave in $x \geq 0$ for any $\sigma \neq 0$. Furthermore, $(1, 0)^T$

and $(0, 1)^T$ are the only global minima for the optimization problem

$$\begin{aligned} \min \quad & \log^{-1} \left(1 + \frac{N}{\sigma^2 + x} \right) + \log^{-1} \left(1 + \frac{N}{\sigma^2 + y} \right) \\ \text{s.t.} \quad & x + y = 1, \quad x \geq 0, \quad y \geq 0, \end{aligned}$$

where $N > 0$.

To establish Claim 1, we first show the strict concavity of $r(x) = \log^{-1} \left(1 + (\sigma^2 + x)^{-1} \right)$. Since

$$\begin{aligned} r'(x) &= \frac{r^2(x)}{(1 + \sigma^2 + x)(\sigma^2 + x)}, \\ r''(x) &= \frac{2r^2(x)(r(x) - (1/2 + \sigma^2 + x))}{(1 + \sigma^2 + x)^2(\sigma^2 + x)^2}, \end{aligned}$$

it follows that

$$\begin{aligned} r''(x) &< 0 \iff \\ g(x) &= \log \left(1 + 1/(\sigma^2 + x) \right) - 1/(1/2 + \sigma^2 + x) > 0. \end{aligned}$$

Let $z = 1/(\sigma^2 + x)$ and consider $h(z) = \log(1 + z) - 2z/(2 + z)$. Since

$$h(0) = 0 \quad \text{and} \quad h'(z) = \frac{z^2}{(1 + z)(z + 2)^2} > 0, \quad \forall z > 0,$$

it follows that $g(x) = h(z) > 0$ for all $x > 0$, implying the strict concavity of $r(x)$ over the interval $(0, \infty)$. Since affine transformation does not change strict concavity of a function, this implies that $\log^{-1} \left(1 + N/(\sigma^2 + x) \right)$ is also strictly concave for $x \geq 0$. Finally, since the minimum of a strictly concave function over a polytope is always attained at a vertex [40], we have established the claim.

To establish Lemma 3.2, let us consider the following parametric optimization problem in \mathbb{R}^2

$$\begin{aligned} \min \quad & \frac{1}{r_4} + \frac{1}{r_3} + \frac{1}{r_2} + \frac{1}{r_1} \\ \text{s.t.} \quad & r_4 = \log \left(1 + 0.9/(1, 1)\mathbf{v}^2 \right), \\ & r_3 = \log \left(1 + 1/(1, 1)\mathbf{v}^2 \right), \\ & r_2 = \log \left(1 + 100/|(1, 0)\mathbf{v}|^2 \right), \\ & r_1 = \log \left(1 + 100/|(0, 1)\mathbf{v}|^2 \right), \\ & \|\mathbf{v}\| = t, \end{aligned} \tag{19}$$

where $\mathbf{v} = (v_1, v_2)^T$, and $t \geq 0$ is a parameter.

Claim 2: Let $f(t)$ denote the minimum value of (19). The following properties hold:

- 1) $f(t)$ is a strictly decreasing function in $[0, 1]$,
- 2) When $t \in (0.95, 1]$, the global minima of (19) are $(t, 0)^T$ and $(0, t)^T$,
- 3) $f'(t)$ is an increasing function in $[0.95, 1]$.

Let us argue that Claim 2 is true. First, if part 2) is true, then we can check part 3) directly by first computing $f(t)$ as the objective value of (19) at its optimal solutions $(t, 0)^T$ or $(0, t)^T$, and then verifying that $f''(t) > 0$ for $t \in [0.95, 1]$. We omit the details of computation for space reason. So we only need to argue parts 1) and 2).

When t is fixed, the KKT condition for problem (19) can be written as

$$\begin{cases} \phi(v_1, v_2) + \psi(v_1) &= \lambda v_1, \\ \phi(v_1, v_2) + \psi(v_2) &= \lambda v_2, \\ v_1^2 + v_2^2 &= t^2, \end{cases} \quad (20)$$

where $\phi(x, y)$ and $\psi(x)$ are given above and λ is the Lagrangian multiplier associated with the constraint $\|\mathbf{v}\| = t$ in (19). It can be seen that $\frac{\sqrt{2}}{2}(t, t)^T$ is always a solution to (20). Moreover, $(t, 0)^T$ and $(0, t)^T$ are the (only) non-differentiable points of (19). Suppose $(\bar{v}_1(t), \bar{v}_2(t))^T$, where $\bar{v}_1(t) \neq \bar{v}_2(t)$, are other KKT solutions to (20) (if there are any). Notice that the global minimum $f(t)$ is always attained at a KKT point or at a non-differentiable point of (19). Thus, we must have

$$f(t) = \min_{\{A, B, C\}} \{f_A(t), f_B(t), f_C(t)\},$$

where $f_A(t)$ and $f_B(t)$ are given in the following page and $f_C(t)$ is the objective value of (19) at $(\bar{v}_1(t), \bar{v}_2(t))^T$.

Next, we claim that $f(t)$ is a decreasing function in $t \in [0, 1]$. We prove this by examining the monotonicity of the component functions $f_A(t)$, $f_B(t)$ and $f_C(t)$. By (20), we have

$$\lambda = \frac{\psi(\bar{v}_1(t)) - \psi(\bar{v}_2(t))}{\bar{v}_1(t) - \bar{v}_2(t)} < 0,$$

where the last step follows from the fact that $\psi(x)$ is a decreasing function for $x \in [0, 1]$. Therefore, we conclude from the standard sensitivity analysis [38] of duality multipliers that $f_C(t)$ decreases as t increases in $[0, 1]$. For $t \in [0, 0.75]$, it can be checked that all of $f_i(t)$, $i = A, B, C$, are decreasing functions. Thus, $f(t)$ decreases monotonically in $[0, 0.75]$. For $t \in (0.75, 0.85)$, the global minimizer of (19) is neither $\frac{\sqrt{2}}{2}(t, t)^T$ nor $(t, 0)^T$ or $(0, t)^T$ as we always can find a point at which the objective function is smaller than $f_A(t)$ and $f_B(t)$. As a result, $f(t) = f_C(t)$ decreases in $t \in (0.75, 0.85)$. For $t \in [0.85, 1]$, we have know $f_B(t) \leq f_A(t)$, so $f(t) = \min\{f_B(t), f_C(t)\}$. As both of $f_B(t)$ and $f_C(t)$ are decreasing functions in $t \in [0.85, 1]$, we know $f(t)$ decreases in $[0.85, 1]$. This completes the proof that $f(t)$ is monotonically decreasing in $[0, 1]$.

We next prove part 2) of Claim 2: namely for $t \in [0.95, 1]$, the minima of (19) are $(t, 0)^T$ and $(0, t)^T$. We can use $v_2 = \sqrt{t^2 - v_1^2}$ to transform (19) into an unconstrained univariate optimization problem, where $-t \leq v_1 \leq t$. The global minimizer $(v_1, v_2)^T$ of (19) should satisfy that $v_1 v_2 \geq 0$, else we can find $(\bar{v}_1, \bar{v}_2)^T$ such that $\bar{v}_1 + \bar{v}_2 = v_1 + v_2$, $|\bar{v}_1| < |v_1|$ and $|\bar{v}_2| < |v_2|$, at which objective would be lower. Thus, we only need to consider the case $v_1 v_2 > 0$. Due to the symmetry of (19), we only consider the case $0 \leq v_1 \leq \sqrt{2}t/2$. It can be checked that for any $t \in [0.95, 1]$, the objective function of (19) increases as v_1 increases in $[0, \sqrt{2}t/2]$. Hence, $(0, t)^T$ and $(t, 0)^T$ are the global minima for (19) with $t \in [0.95, 1]$.

Claim 3: When $N \geq 2(e^{40m} - 1)$, the global minima \mathbf{v}_{4i}^* of (8) must have unit-norm $\|\mathbf{v}_{4i}^*\| = 1, i = 1, 2, \dots, n$.

By symmetry, we only need to prove $\|\mathbf{v}_4^*\| = 1$. Let us consider the following parametric optimization problem (21) in \mathbf{v}_4 with the other variables $\mathbf{v}_{4i}, i = 2, 3, \dots, n$, fixed, and t as a parameter. Let $g(t)$ be the optimum value of (21) and let $g_1(t), g_2(t)$ denote respectively the values of

$$\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} + \frac{1}{r_4} \quad \text{and} \quad \sum_{j=1}^m \left(\frac{1}{r_{4n+2j}} + \frac{1}{r_{4n+2j-1}} \right)$$

when evaluated at the global minima of (21). By definition, $f(t) \leq g_1(t)$ for all $t \in [0, 1]$. We claim that $f(t) = g_1(t)$ for $t \in [0.95, 1]$. To see this, we consider the optimization problem (22). Define $|\mathbf{h}_a^\dagger \mathbf{v}_4|^2 = x$, then $|\mathbf{h}_b^\dagger \mathbf{v}_4|^2 = t^2 - x$, and the objective function of (22) is only dependent on x , while its constraint $\|\mathbf{v}_4\| = t$ can be transformed to $0 \leq x \leq t^2$. Using Claim 1, we see that the objective function of (22) is strictly concave in x . As a result, the globally optimal x can only be attained at the end points of the constraint interval $[0, t^2]$. In other words, the global minimizer of (22) must be either $(t, 0)^T$ or $(0, t)^T$. Since the global minima of (19) are also $(t, 0)^T$ and $(0, t)^T$ when $t \in [0.95, 1]$, it follows that the global minimizer of (21) is either $(t, 0)^T$ or $(0, t)^T$, if $t \in [0.95, 1]$. This implies that $f(t) = g_1(t)$ in $[0.95, 1]$.

Next, we prove that if $N \geq 2(e^{40m} - 1)$, then $g(1) < g(t)$ for $t \in [0, 1]$. Due to the choice of N , it can be checked that

$$f'(1) + \frac{2m}{\log^2(1 + N/3)} \leq 0,$$

$$f(0.95) \geq f(1) + \frac{2m}{\log(1 + N/2)}.$$

Then, we have

1) For any $t \in [0.95, 1]$,

$$\begin{aligned} g'(t) &= g'_1(t) + g'_2(t) = f'(t) + g'_2(t) \\ &\leq f'(1) + g'_2(t) \\ &< f'(1) + \frac{2m}{\log^2(1 + N/3)} \leq 0. \end{aligned}$$

The second equality is due to the fact $f(t) = g_1(t)$ for $t \in [0.95, 1]$. The first inequality holds since $f'(t)$ is an increasing function in $[0.95, 1]$ (cf. part 3) of Claim 2), while the second inequality follows from

$$g'_2(t) < \frac{2m}{\log^2(1 + N/3)}, \quad \text{for all } t \in [0.95, 1].$$

The non-positiveness of $g'(t)$ over $t \in [0.95, 1]$ implies $g(t) > g(1)$, for $t \in [0.95, 1]$.

2) For any $t \in [0, 0.95]$, we have

$$\begin{aligned} g(t) &> g_1(t) \geq f(t) \\ &\geq f(0.95) \geq f(1) + \frac{2m}{\log(1 + N/2)} \\ &\geq g_1(1) + g_2(t) = g(1), \end{aligned}$$

where the third inequality holds because $f(t)$ is a strictly decreasing function in $[0, 1]$ (cf. part 1) of Claim 2) and the last inequality is due to

$$\frac{2m}{\log(1 + N/2)} \geq g_2(t), \quad \text{for all } t \in [0, 1].$$

$$\begin{aligned}\phi(x, y) &= \frac{-1.8(x+y)}{\log^2\left(1+0.9(x+y)^2\right)\left(1+0.9(x+y)^2\right)} + \frac{2}{\log^2\left(1+1/(x+y)^2\right)\left((x+y)^2+1\right)(x+y)}, \\ \psi(x) &= \frac{200}{\log^2(1+100/x^2)(100+x^2)x},\end{aligned}$$

$$\begin{cases} f_A(t) &= \frac{1}{\log(1+1.8t^2)} + \frac{1}{\log(1+0.5/t^2)} + \frac{2}{\log(1+200/t^2)}, \\ f_B(t) &= \frac{1}{\log(1+0.9t^2)} + \frac{1}{\log(1+1/t^2)} + \frac{1}{\log(1+100/t^2)}, \end{cases}$$

$$\begin{aligned}\min_{\mathbf{v}_4} \quad & \left(\frac{1}{r_4} + \frac{1}{r_3} + \frac{1}{r_2} + \frac{1}{r_1} \right) + \sum_{j=1}^m \left(\frac{1}{r_{4n+2j}} + \frac{1}{r_{4n+2j-1}} \right) \\ \text{s.t.} \quad & r_4 = \log\left(1+0.9|(1,1)\mathbf{v}_4|^2\right), \\ & r_3 = \log\left(1+1/|(1,1)\mathbf{v}_4|^2\right), \\ & r_2 = \log\left(1+100/|(1,0)\mathbf{v}_4|^2\right), \\ & r_1 = \log\left(1+100/|(0,1)\mathbf{v}_4|^2\right), \\ & \|\mathbf{v}_4\| = t, \\ & r_{4n+2j} = \log\left(1+N/\left(1+|\mathbf{h}_{\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2\right)\right), \\ & r_{4n+2j-1} = \log\left(1+N/\left(1+|\mathbf{h}_{-\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{-\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2\right)\right), \\ & 1 \leq j \leq m.\end{aligned}\tag{21}$$

$$\begin{aligned}\min_{\mathbf{v}_4} \quad & \sum_{j=1}^m \left(\frac{1}{r_{4n+2j}} + \frac{1}{r_{4n+2j-1}} \right) \\ \text{s.t.} \quad & \|\mathbf{v}_4\| = t, \\ & r_{4n+2j} = \log\left(1+N/\left(1+|\mathbf{h}_{\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2\right)\right), \\ & r_{4n+2j-1} = \log\left(1+N/\left(1+|\mathbf{h}_{-\pi(j)}^\dagger \mathbf{v}_{4|\pi(j)|}|^2 + |\mathbf{h}_{-\tau(j)}^\dagger \mathbf{v}_{4|\tau(j)|}|^2\right)\right), \\ & 1 \leq j \leq m.\end{aligned}\tag{22}$$

Combining the above two steps shows $g(t) > g(1)$ for all $t \in [0, 1)$. It follows that the minimum of $g(t)$ over $[0, 1]$ is attained at $t = 1$. This establishes Claim 3.

Finally, notice that Claim 3 implies that problem (8) is equivalent to problem (21) with $t = 1$, when $N \geq 2(e^{40m} - 1)$. Since the global optimal solution of (21) must be either $(0, t)^T$ or $(t, 0)^T$ for $t \in [0.95, 1]$, it follows that the optimal solutions $\mathbf{v}_{4i}^*, i = 1, 2, \dots, K$, for problem (8) should be either $(1, 0)^T$ or $(0, 1)^T$ when $N \geq 2(e^{40m} - 1)$.

APPENDIX II PROOF OF LEMMA 3.3

Proof: Let $r(x) = \log(1+1/(b+x))$, then we have $f(x) = \log r(x)$, and

$$r'(x) = \frac{-1}{(b+1+x)(b+x)}, \quad f'(x) = \frac{r'(x)}{r(x)}$$

and

$$f''(x) = \frac{(g(x) - 1)}{(b+1+x)^2(b+x)^2 r^2(x)},$$

where $g(x) = (2b+2x+1)r(x)$ and $b = \sigma^2 \geq 0$. Let $y = x+b$, then $g(x)$ becomes $h(y) = (2y+1)\log(1+1/y)$. It suffices to prove that $h(y) \geq 1$ for all $y \geq 0$. Since

$$h'(y) = 2\log\left(1 + \frac{1}{y}\right) - \frac{(2y+1)}{y(y+1)}$$

and

$$h''(y) = \frac{1}{y^2(y+1)^2} > 0,$$

we know $h'(y)$ is an increasing function. Since $\lim_{y \rightarrow +\infty} h'(y) = 0$, it follows that $h'(y) \leq 0$ and $h(y)$ is a decreasing function. Notice that $\lim_{y \rightarrow +\infty} h(y) = 2$. Thus, we have $h(y) \geq 2 > 1$ for all $y \geq 0$. This further implies that $\log \log(1+1/(b+x))$ is strictly convex for $x \geq 0$. ■

APPENDIX III CONVERGENCE OF THE CYCLIC COORDINATE DESCENT ALGORITHM

We first need to estimate the step length.

Claim 1: Suppose $c_1 \in (0, 1)$, $\nabla h(\mathbf{y}^i)^T \mathbf{d}^i < 0$ and $\max_{\mathbf{y} \in Y} \|\nabla^2 h(\mathbf{y})\| \leq B$, where $h(\mathbf{y})$ is a multi-variable function associated with \mathbf{y} and Y is the feasible region. Consider the Armijo step length rule [38] whereby $\alpha_i := \gamma^\ell$, with $\gamma \in (0, 1)$ and $\ell \geq 0$ being the smallest integer satisfying

$$h(\mathbf{y}^i + \gamma^\ell \mathbf{d}^i) \leq h(\mathbf{y}^i) + c_1 \gamma^\ell \nabla h(\mathbf{y}^i)^T \mathbf{d}^i. \quad (23)$$

Then we have

$$1 \geq \alpha_i \geq \min \left\{ 1, \frac{2(c_1 - 1) \nabla h(\mathbf{y}^i)^T \mathbf{d}^i}{B \gamma \|\mathbf{d}^i\|^2} \right\}, \quad (24)$$

and

$$\begin{aligned} & h(\mathbf{y}^i) - h(\mathbf{y}^i + \alpha_i \mathbf{d}^i) \\ & \geq \min \left\{ -c_1 \nabla h(\mathbf{y}^i)^T \mathbf{d}^i, \frac{2c_1(1 - c_1) (\nabla h(\mathbf{x}^i)^T \mathbf{d}^i)^2}{B \gamma \|\mathbf{d}^i\|^2} \right\}. \end{aligned} \quad (25)$$

Let us argue Claim 1 holds. Suppose that the step length $\alpha_i = 1$ is not accepted. In this case, α_i will be the largest step satisfying the sufficient decrease condition (23), implying

$$h(\mathbf{y}^i + \alpha_i \gamma^{-1} \mathbf{d}^i) > h(\mathbf{y}^i) + c_1 \alpha_i \gamma^{-1} \nabla h(\mathbf{y}^i)^T \mathbf{d}^i. \quad (26)$$

By Taylor expansion, (27) in the next page holds true. Combining (26) and (27) yields (24). Substituting (24) into the sufficient decrease condition (23), we immediately obtain (25), which establishes Claim 1.

We now proceed with the proof of Theorem 4.1. The basic idea is based on the contradiction principle. More exactly, if there is no convergence, we can find one descent direction which can provide a sufficient descent in the objective function value and then obtain a contradiction.

At first, since the iterates $\{\mathbf{x}^i\}$ lie in a compact set X , there must exist an accumulation point for $\{\mathbf{x}^i\}$. Let $\bar{\mathbf{x}}$ denote an accumulation point such that

$$\bar{\mathbf{x}} = \lim_{i \in \mathcal{I}_0, i \rightarrow \infty} \mathbf{x}^i$$

for some subsequence indexed by \mathcal{I}_0 . Since the feasible set X is closed, $\bar{\mathbf{x}}$ must also be feasible. Furthermore, since the projection mapping and the function f are both continuous, it follows that

$$\begin{aligned} \lim_{i \in \mathcal{I}_0, i \rightarrow \infty} \mathbf{d}_1^{i+1} &= \lim_{i \in \mathcal{I}_0, i \rightarrow \infty} P_{X_1}(\mathbf{x}_1^i - \nabla_{\mathbf{x}_1} f(\mathbf{z}_0^i)) - \mathbf{x}_1^i \\ &= P_{X_1}[\bar{\mathbf{x}}_1 - \nabla_{X_1} f(\bar{\mathbf{x}})] - \bar{\mathbf{x}}_1 \triangleq \bar{\mathbf{d}}_1 \end{aligned}$$

and

$$\lim_{i \in \mathcal{I}_0, i \rightarrow \infty} f(\mathbf{x}^i) = f(\bar{\mathbf{x}}).$$

Notice that the function values $\{f(\mathbf{x}^i)\}$ are decreasing and bounded below, then $f(\mathbf{x}^i) \rightarrow f(\bar{\mathbf{x}})$. Since $f(\mathbf{x})$ is twice continuously differentiable and the feasible set X is bounded, it follows that

$$B = \max_{k=1,2,\dots,K} \max_{\mathbf{x}_k \in X_k} \|\nabla_{\mathbf{x}_k}^2 f(\mathbf{x})\| < \infty.$$

Because the projection operator is non-expansive and $\nabla f(\mathbf{x})$ is continuous in bounded region, we obtain from (17) that

$$\begin{aligned} \|\mathbf{d}_k^{i+1}\| &= \|P_{X_k}[\mathbf{x}_k^i - \nabla_{\mathbf{x}_k} f(\mathbf{z}_{k-1}^{i+1})] - P_{X_k}[\mathbf{x}_k^i]\| \\ &\leq \|[\mathbf{x}_k^i - \nabla_{\mathbf{x}_k} f(\mathbf{z}_{k-1}^{i+1})] - \mathbf{x}_k^i\| = \|\nabla_{\mathbf{x}_k} f(\mathbf{z}_{k-1}^{i+1})\|. \end{aligned}$$

Denoting $\max_{\mathbf{x} \in X} \|\nabla f(\mathbf{x})\| \triangleq M$, hence we have

$$\|\mathbf{d}_k^{i+1}\| \leq \|\nabla_{\mathbf{x}_k} f(\mathbf{z}_{k-1}^{i+1})\| \leq M < +\infty. \quad (28)$$

The lower bound and upper bound on $\|\nabla_{\mathbf{x}_k} f(\mathbf{z}_{k-1}^{i+1})\|$ is very helpful in estimating the decrease value of the function.

We proceed by contradiction and suppose $\bar{\mathbf{x}}$ is not a KKT point. Then $\bar{\mathbf{d}} = P_X[\bar{\mathbf{x}} - \nabla f(\bar{\mathbf{x}})] - \bar{\mathbf{x}} \neq 0$ so that $\delta = \|\bar{\mathbf{d}}\| > 0$. Let

$$k^* = \min\{k \mid \delta_k = \|\bar{\mathbf{d}}_k\| > 0\},$$

and suppose $k^* > 1$ without loss of generality. By definition, we have

$$\lim_{i \in \mathcal{I}_0, i \rightarrow \infty} \|\mathbf{d}_k^{i+1}\| = \delta_k = 0, \quad \text{for } k < k^*.$$

Recall the definition of \mathbf{z}_k^{i+1} (16). Since

$$\begin{aligned} \|\mathbf{z}_1^{i+1} - \bar{\mathbf{x}}\| &= \|\mathbf{x}_1^i + \alpha_1^{i+1} \mathbf{d}_1^{i+1} - \bar{\mathbf{x}}_1\| \\ &\leq \|\mathbf{x}_1^i - \bar{\mathbf{x}}_1\| + \alpha_1^{i+1} \|\mathbf{d}_1^{i+1}\| \\ &\leq \|\mathbf{x}_1^i - \bar{\mathbf{x}}_1\| + \|\mathbf{d}_1^{i+1}\| \xrightarrow{i \in \mathcal{I}_0, i \rightarrow \infty} 0, \end{aligned}$$

we have $\lim_{i \in \mathcal{I}_0, i \rightarrow \infty} \mathbf{z}_1^{i+1} = \bar{\mathbf{x}}$. In general, the same argument shows that

$$\lim_{i \in \mathcal{I}_0, i \rightarrow \infty} \mathbf{z}_k^{i+1} = \bar{\mathbf{x}}, \quad \forall k < k^*.$$

Consequently, there holds

$$\begin{aligned} \lim_{i \in \mathcal{I}_0, i \rightarrow \infty} \nabla_{\mathbf{x}_{k^*}} f(\mathbf{z}_{k^*-1}^{i+1}) &= \nabla_{\mathbf{x}_{k^*}} f(\bar{\mathbf{x}}) \\ \lim_{i \in \mathcal{I}_0, i \rightarrow \infty} \mathbf{d}_{k^*}^{i+1} &= \bar{\mathbf{d}}_{k^*}. \end{aligned} \quad (29)$$

Let us use $\bar{\theta}_{k^*}$ to denote the angle between $\bar{\mathbf{d}}_{k^*}$ and $\nabla_{\mathbf{x}_{k^*}} f(\bar{\mathbf{x}})$. Since $\bar{\mathbf{y}}_{k^*} = \bar{\mathbf{x}}_{k^*} - \nabla_{\mathbf{x}_{k^*}} f(\bar{\mathbf{x}}) \notin X_{k^*}$, it follows from the property of projection that

$$(\bar{\mathbf{x}}_{k^*} - P_{X_{k^*}}[\bar{\mathbf{y}}_{k^*}])^T (\bar{\mathbf{y}}_{k^*} - P_{X_{k^*}}[\bar{\mathbf{y}}_{k^*}]) \leq 0,$$

which further implies

$$\begin{aligned} & \|P_{X_{k^*}}[\bar{\mathbf{y}}_{k^*}] - \bar{\mathbf{x}}_{k^*}\|^2 \\ & \leq (\bar{\mathbf{x}}_{k^*} - P_{X_{k^*}}[\bar{\mathbf{y}}_{k^*}])^T (\bar{\mathbf{x}}_{k^*} - \bar{\mathbf{y}}_{k^*}) \\ & = \|P_{X_{k^*}}[\bar{\mathbf{y}}_{k^*}] - \bar{\mathbf{x}}_{k^*}\| \|\nabla_{\mathbf{x}_{k^*}} f(\bar{\mathbf{x}})\| \cos \bar{\theta}_{k^*}. \end{aligned}$$

Canceling the factor $\|P_{X_{k^*}}[\bar{\mathbf{y}}_{k^*}] - \bar{\mathbf{x}}_{k^*}\| \neq 0$ from both sides yields

$$\begin{aligned} \delta_{k^*} = \|\bar{\mathbf{d}}_{k^*}\| &= \|P_{X_{k^*}}[\bar{\mathbf{y}}_{k^*}] - \bar{\mathbf{x}}_{k^*}\| \\ &\leq \|\nabla_{\mathbf{x}_{k^*}} f(\bar{\mathbf{x}})\| \cos \bar{\theta}_{k^*} \\ &\leq M \cos \bar{\theta}_{k^*}, \end{aligned}$$

where the last inequality follows from (28). Thus, we have

$$\cos \bar{\theta}_{k^*} \geq \frac{\delta_{k^*}}{M}. \quad (30)$$

Now we can use the property (28), (29) and (30) to conclude

$$\|\nabla_{\mathbf{x}_{k^*}} f(\mathbf{z}_{k^*-1}^{i+1})\| \geq \|\mathbf{d}_{k^*}^{i+1}\| \geq \frac{\delta_{k^*}}{2} \quad \text{and} \quad \cos \theta_{k^*}^{i+1} \geq \frac{\delta_{k^*}}{2M},$$

for all $i \in \mathcal{I}_0$ and $i \geq i_0$, where i_0 is a sufficiently large integer and $\theta_{k^*}^{i+1}$ denotes the angle between $\nabla_{\mathbf{x}_{k^*}} f(\mathbf{z}_{k^*-1}^{i+1})$ and $\mathbf{d}_{k^*}^{i+1}$.

Now we can use (23)-(25) of Claim 1 to obtain a contradiction. In particular, we consider (31), which is a contradiction. The fourth inequality is due to the sufficient decrease condition and the fifth inequality is due to Claim 1. Therefore, $\bar{\mathbf{x}}$ is a stationary point.

$$\begin{aligned}
h(\mathbf{y}^i + \alpha_i \gamma^{-1} \mathbf{d}^i) &= h(\mathbf{y}^i) + \alpha_i \gamma^{-1} \nabla h(\mathbf{y}^i)^T \mathbf{d}^i + \frac{\alpha_i^2 \gamma^{-2}}{2} \mathbf{d}^{iT} \nabla^2 h(\xi) \mathbf{d}^i \\
&\leq h(\mathbf{y}^i) + \alpha_i \gamma^{-1} \nabla h(\mathbf{d})^T \mathbf{d}^i + \frac{\alpha_i^2 \gamma^{-2}}{2} B \|\mathbf{d}^i\|^2.
\end{aligned} \tag{27}$$

$$\begin{aligned}
+\infty &> \sum_i (f(\mathbf{x}^i) - f(\mathbf{x}^{i+1})) \geq \sum_{i \in \mathcal{I}_0} (f(\mathbf{x}^i) - f(\mathbf{x}^{i+1})) \\
&\geq \sum_{i \in \mathcal{I}_0} (f(\mathbf{z}_{k^*-1}^{i+1}) - f(\mathbf{z}_{k^*}^{i+1})) \geq \sum_{i \in \mathcal{I}_0} \left(-c_1 \alpha_{k^*}^{i+1} \nabla_{\mathbf{x}_{k^*}} f(\mathbf{z}_{k^*-1}^{i+1})^T \mathbf{d}_{k^*}^{i+1} \right) \\
&\geq \sum_{i \in \mathcal{I}_0} \left(\frac{2c_1(1-c_1) \left(\nabla_{\mathbf{x}_{k^*}} f(\mathbf{z}_{k^*-1}^{i+1})^T \mathbf{d}_{k^*}^{i+1} \right)^2}{B\gamma \|\mathbf{d}_{k^*}^{i+1}\|^2} \right) \\
&= \sum_{i \in \mathcal{I}_0} \left(\frac{2c_1(1-c_1) \left(\|\nabla_{\mathbf{x}_{k^*}} f(\mathbf{z}_{k^*-1}^{i+1})\| \|\mathbf{d}_{k^*}^{i+1}\| \cos \theta_{k^*}^{i+1} \right)^2}{B\gamma \|\mathbf{d}_{k^*}^{i+1}\|^2} \right) \\
&= \sum_{i \in \mathcal{I}_0} \left(\frac{2c_1(1-c_1) \left(\|\nabla_{\mathbf{x}_{k^*}} f(\mathbf{z}_{k^*-1}^{i+1})\| \cos \theta_{k^*}^{i+1} \right)^2}{B\gamma} \right) \\
&\geq \sum_{i \geq i_0, i \in \mathcal{I}_0} \left(\frac{c_1(1-c_1) \delta_{k^*}^4}{8B\gamma M^2} \right) = +\infty,
\end{aligned} \tag{31}$$

REFERENCES

- [1] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, Special Issue on Signal Processing and Networking for Dynamic Spectrum Access, vol. 2, no. 1, pp. 57-73, Feb. 2008.
- [2] A. Wiesel, Y. C. Eldar, and S. S. Shitz, "Linear precoding via conic programming for fixed MIMO receivers," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 161-176, Jan. 2006.
- [3] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmitter beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437-1450, Oct. 1998.
- [4] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *Proc. IEEE Vehicular Technology Conf. (VTC)*, pp. 87-91, Sep. 2002.
- [5] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," accepted in *IEEE Transactions on Wireless Communications*, 2010.
- [6] Z.-Q. Luo and S. Zhang, "Duality gap estimation and polynomial time approximation for optimal spectrum management," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2675-2689, Jul. 2009.
- [7] R. Zakhour and D. Gesbert, "Coordination on the MISO interference channel using the virtual SINR framework," in *Proc. WSA'09*, 2009.
- [8] E. A. Jorswieck, E. G. Larsson and D. Danev, "Complete characterization of the pareto boundary for the MISO interference channel," in *Proc. IEEE Trans. Signal Processing*, Aug. 2008.
- [9] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, submitted for publication.
- [10] O. Somekh, O. Simeone, Y. Bar-Ness, and A. Haimovich, "Distributed multi-cell zero-forcing beamforming in cellular downlink channels," *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, pp. 1-6, 2006.
- [11] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 56-61, Aug. 2006.
- [12] G. Foschini, M. Karakayali, and R. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," *IEEE Proc. Commun.*, vol. 153, no. 4, pp. 548-555, Aug. 2006.
- [13] B. O. Lee, H. W. Je, I. Sohn, O.-S. Shin, and K. B. Lee, "Interference-aware Decentralized Precoding for Multicell MIMO TDD Systems," in *Proc. IEEE GLOBECOM08*, 2008.
- [14] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Distributed multicell and multiantenna precoding: characterization and performance evaluation," in *Proc. IEEE GLOBECOM09*, 2009.
- [15] E. Larsson and E. Jorswieck, "Competition versus Cooperation On The Miso Interference Channel," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1059-1069, 2008.
- [16] B. L. Ng, J. Evans, S. Hanly, and D. Aktas, "Distributed Downlink Beamforming with Cooperative Base Stations," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5491-5499, 2008.
- [17] D. Schmidt, C. Shi, R. Berry, M. Honig, and W. Utschick, "Distributed resource allocation schemes: Pricing algorithms for power control and beamformer design in interference networks," *IEEE Signal Processing Magazine*, special issue on "Game Theory in Signal Processing and Communications", vol. 26, no. 5, pp. 53-63, Sep. 2009.
- [18] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1912-1921, 2003.
- [19] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 361-374, 2006.
- [20] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646-2660, 2007.
- [21] P. Marsch and G. Fettweis, "On multicell cooperative transmission in backhaul-constrained cellular systems," *Ann. Telecommun.*, vol. 63, pp. 253-269, 2008.
- [22] S. Jing, D. Tse, J. Soriaga, J. Hou, J. Smee, and R. Padovani, "Multicell downlink capacity with coordinated processing," *EURASIP J. Wirel. Commun. Netw.*, 2008.
- [23] X. Shang, B. Chen and H.V. Poor, "Multi-User MISO Interference Channels with Single-User Detection: Optimality of Beamforming and the Achievable Rate Region," submitted to *IEEE trans. Inf. Theory*. <http://arxiv.org/abs/0907.0505>, Apr. 2009.
- [24] R. Mochaourab and E. Jorswieck, "Optimal Beamforming in Interference Networks with Perfect Local Channel Information," submitted to *IEEE Trans. on Signal Processing*. <http://arxiv.org/abs/1004.4492>, Apr. 2010.
- [25] V. S. Annapureddy and V. V. Veeravalli, "Sum Capacity of MIMO Interference Channels in the Low Interference Regime," submitted to *IEEE trans. Inf. Theory*. <http://arxiv.org/abs/0909.2074>, Sep. 2009.
- [26] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, San Francisco, USA: W. H. Freeman and Company, 1979.
- [27] X. Shang, B. Chen, G. Kramer and H.V. Poor, "Noisy-interference sum-rate capacity of parallel Gaussian interference channel," submitted to *IEEE trans. on Inf. Theory*. <http://arxiv.org/abs/0903.0595>, Mar. 2009.

- [28] C. W. Tan, M. Chiang, R. Srikant, "Maximizing sum rate and minimizing MSE on multiuser downlink: optimality, fast algorithms and equivalence via max-min SIR," in *Proc. ISIT*, 2009.
- [29] S. Friedland and C. W. Tan, "Maximizing sum rates in Gaussian interference-limited channels," submitted to *Numerical Analysis*. <http://arxiv.org/abs/0806.2860v2>, Jul. 2008.
- [30] C. H. Papadimitriou, *Computational Complexity*, Beijing, China: Tsinghua University Press, 2004.
- [31] E. Björnson and B. Ottersten, "A framework for trained-based estimation in arbitrarily correlated Rician MIMO channels with Rician ditribubance," *IEEE Trans. Signal Process.*, submitted for publication.
- [32] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distribution Computation: Numerical Methods*. Prentice-Hall, Inc., Englewoodd Cliffs, NJ, 1989.
- [33] M. J. D. Powell, "On search direction for minimization algorithms," *Mathematical Programming* 4, pp. 193-201, 1973.
- [34] J. C. Bezdek, R. J. Hathaway, R. E. Howard, C. A. Wilson and M. P. Windham, "Local convergence analysis of a grouped variable version of coordinate descent," *Journal of Optimzation Theory and Applications*, vol. 54, no. 3, Sep. 1987.
- [35] M. V. Solodov, "On the convergence of constrained parallel variable distribution algorithm," *SIAM J. OPTM*, vol. 8, no. 1, pp. 187-196, Feb. 1998.
- [36] Z.-Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *Journal of Operation Theory and Applications*, vol. 72, no. 1, pp. 7-35, 1992.
- [37] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Infor. Theory*, vol. 18, no. 4, pp. 460-473, Jul. 1972.
- [38] J. Nocedal and S. J. Wright, *Numerical Optimization*. Germany, Springer Science+Business Media, Inc, 1999.
- [39] Y.-H. Dai and L.-Z. Liao, " R -linear convergence of the Barzilai and Borwein gradient method," *IMA Journal of Numerical Analysis* 22, pp. 1-10, 2002.
- [40] H. L. Hoffman, "A method for globally minimizing concave functions over convex sets," *Mathematical Programming*, vol. 20, no. 1, pp. 22-32, 1981.